

Quantum filtering using POVM measurements: Conditional Expectation

Ram A. Somaraju, Alain Sarlette and Hugo Thienpont

Abstract—The objective of this work is to develop a recursive, discrete time quantum filtering equation for a system that interacts with a probe, on which measurements are performed according to the Positive Operator Valued Measures (POVMs) framework. POVMs are the most general measurements one can make on a quantum system and although in principle they can be reformulated as projective measurements on larger spaces, for which filtering results exist, a direct treatment of POVMs can be more natural and simplify the filter computations for some applications. Hence we formalize the notion of commuting (Davies) instruments which allows one to develop joint measurement statistics of two POVM type measurements as explicit functions of the POVMs. This allows us to prove the existence of a notion of conditional expectation POVM, which is essential for the development of a filtering equation. We demonstrate that under generally satisfied assumptions, the reduced model given by POVM elements is sufficient for the purpose of our quantum filtering task.

I. INTRODUCTION

The theory of filtering considers the estimation of the system state from noisy and/or partial observations (see e.g. [1]). For quantum systems, filtering theory was initiated in the 1980s by Belavkin in a series of papers [2], [3], [4], [5]. Belavkin makes use of the operational formalism of Davies [6], which is a precursor to the theory of quantum filtering. He has also realized that due to the unavoidable back-action of quantum measurements, the theory of filtering plays a fundamental role in quantum feedback control (see e.g. [3], [5]). The theory of quantum filtering was independently developed in the physics community, particularly in the context of quantum optics, under the name of quantum trajectory theory [7], [8], [9], [10].

The basic model used to derive filtering equations for a quantum system uses a system-probe interaction. A quantum system, whose state needs to be estimated, is made to interact with a probe and the state of the system becomes entangled with that of the probe. After this interaction, an observable is measured on the probe and this measurement outcome is used to estimate the state of the system. The commutativity of

any system observable with any probe observable is used to develop a recursive Markov filtering equation for the system observables (see e.g. [11], [12] for an excellent tutorial review of these ideas).

Suppose \mathcal{H}_S is the Hilbert space corresponding to the system whose state needs to be estimated and \mathcal{H}_P is the Hilbert space of the probe. According to the classical von Neumann definition, any probe observable is a self-adjoint operator Q on \mathcal{H}_P ; the measurement of such an observable results in an outcome that is (stochastically) an eigenvalue of Q and the probe state after measurement gets projected onto the corresponding eigenspace of Q . As far as we are aware, all discussions on quantum filtering theory so far have assumed that the probe undergoes such a von Neumann measurement, also called projective measurement or Projection Valued Measure (PVM). However, a more modern treatment of quantum measurement theory shows that the most general possible quantum measurements that one can perform are the so-called Positive Operator Valued Measures¹ (POVMs), of which von Neumann measurements are merely a special case, where all the positive operators in the POVMs are commuting projections [6], [13]. See Section II for a brief overview of POVMs.

POVMs on \mathcal{H}_P can be reformulated as the restriction to \mathcal{H}_P of a PVM on a larger space. However, there is no canonical PVM that corresponds uniquely to a given POVM. This is closely related to the fact that the state of a quantum system after a POVM measurement is not uniquely determined as a function of the POVM. To remedy the latter situation, Davies [6, Ch.3] has shown that one can associate a (non-unique) instrument to any POVM, which determines a completely positive map that specifies the state after measurement conditioned on the measurement outcome. However, there is again no canonical instrument that corresponds to a given POVM. Therefore, it is impossible to uniquely specify the post-measurement state for a given POVM measurement outcome unless the instrument associated with the measurement is known. As stated by Nielsen and Chuang [14, P. 91], “POVMs are best viewed as [...] providing the simplest means by which one can study general measurement statistics, without the necessity for knowing the post-measurement state.” As such they are a minimal description of quantum measurements, so one can hope that the POVM formalism leads to more concise and fast filtering equations, suited for (possibly analog) implementation in real-time quantum feedback experiments.

There are other reasons to develop a POVM-based filtering

Ram A. Somaraju and Hugo Thienpont are with the Brussels Photonics Team, Dept. of Applied Physics and Photonics, Vrije Universiteit Brussel, Pleinlaan 2-1050 Brussels, Belgium. Alain Sarlette is with the SYSTeMS research group, Faculty of Engineering and Architecture, Ghent University, Technologiepark Zwijnaarde 914, 9052 Zwijnaarde (Ghent), Belgium. a.somaraju@gmail.com, alain.sarlette@ugent.be, hthienpo@b-phot.org

R. A. Somaraju and H. Thienpont would like to thank Methusalem project IPARC@VUB for financial support as well as the BELSPO IAP project Photonics@BE. A. Sarlette and R. A. Somaraju are members of the BELSPO IAP project DYSCO.

This paper presents research results funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office. The scientific responsibility rests with its authors.

¹The terminology for such measurements is not standard and other terms such as Positive Operator Measures or generalized measurements are also used.

theory that shortcuts the lift to PVMs in larger spaces. For one, once the POVM theory is available it can be more natural to use, as several practical measurement setups are based on POVMs, such as e.g. approximate position/momentum measurements [6, Ch. 3] or phase measurements [13], [15]. In some infinite-dimensional situations there can even be conceptual barriers to a PVM viewpoint. Indeed, with phase measurements after much investigation there is still no universal agreement on an acceptable PVM [16]. Finally, regarding system identification, the POVM associated with any experimental setup can (at least conceptually) be directly deduced from measurement outcomes; in contrast, in order to ascertain the associated instrument it is necessary to analyze the post-measurement state of the system (for more details see Section II-A). This is not generally feasible in practical experimental setups where often the measurements destroy the quantum state (e.g. photo-detection) and/or non-measurably alter it due to interaction with the environment.

In this paper, we develop a discrete time filtering equation for the system state conditioned on POVM measurements performed on a probe. After reviewing the POVM formalism (Section II), we provide a general theory about the commuting instruments that are associated to POVMs (Section III). This allows us to define conditional expectations POVMs in the measurement-motivated spirit of [11], [12] (Section IV). In the setup consisting of a probe coupled to a target system, any (physically reasonable) instrument associated with a POVM acting only on the probe commutes with any instrument associated with a POVM acting only on the target system. In Section V we show how this allows to define a filtering equation for the system state conditioned on POVM measurement outcomes. A notable result is that this filtering equation is only a function of the probe POVM and does not depend on the associated instrument.

II. REVIEW: POVMS AND ASSOCIATED INSTRUMENTS

The POVM formalism is a standard part of most modern quantum information textbooks. We briefly review it in this section and refer the interested reader to [6, Ch. 3] for more details.

Consider a quantum system with Hilbert space \mathcal{H} , i.e. the system state is given by a density operator ρ which is a unit-trace nonnegative self-adjoint linear operator on \mathcal{H} . We use $*$ to denote the adjoint. Denote by $\mathcal{L}(\mathcal{H})$ the set of linear operators on \mathcal{H} , by $\mathcal{L}_+(\mathcal{H}) \subset \mathcal{L}(\mathcal{H})$ the set of self-adjoint nonnegative linear operators, and by $\mathcal{S}(\mathcal{H}) \subset \mathcal{L}_+(\mathcal{H})$ the set of all possible density operators. Standard textbook treatment of quantum measurements assumes that any physically measurable quantity \hat{A} is associated to a self-adjoint operator $A : \mathcal{H} \rightarrow \mathcal{H}$. Because A is self-adjoint, we have the spectral decomposition²

$$A = \sum_{\omega \in \Omega} \omega P_\omega \quad (1)$$

where Ω is the set of eigenvalues of A and P_ω is the eigenprojection corresponding to eigenvalue ω . Starting with a

system in state ρ , according to von Neumann's measurement postulates we have:

- 1) any measurement of the observable A gives some outcome $\omega \in \Omega$ with probability $\text{Tr}\{\rho P_\omega\}$, and
- 2) after measurement outcome ω , the state of the system becomes

$$\rho' = \frac{P_\omega \rho P_\omega}{\text{Tr}\{P_\omega \rho P_\omega\}}.$$

The first postulate can be thought of as follows: the set Ω is the set of all possible measurement outcomes of an experimental setup (\hat{A}) and to each $\omega \in \Omega$, one assigns a projection P_ω in the Hilbert space \mathcal{H} such that $\text{Tr}\{\rho P_\omega\}$ is the probability of measuring ω . This motivates the following generalization of quantum observables.

Definition 2.1: [6, Def 3.1.1] Let Ω be a set, \mathcal{F} a σ -field of subsets of Ω , and \mathcal{H} a Hilbert space. Then a \mathcal{H} -valued Positive Operator Valued Measure (POVM) on Ω is a map $\hat{A} : \mathcal{F} \rightarrow \mathcal{L}_+(\mathcal{H})$; $E \rightarrow \hat{A}(E) = \hat{A}_E$ such that

- 1) $\hat{A}(E) \geq \hat{A}(\emptyset) = 0$ for all $E \in \mathcal{F}$;
- 2) For any countable, mutually disjoint collection $\{E_n\} \subset \mathcal{F}$ we have

$$\hat{A}\left(\bigcup_n E_n\right) = \sum_n \hat{A}(E_n)$$

where the series convergence on the right is in the weak operator topology;

- 3) $\hat{A}(\Omega) = \mathbb{I}_{\mathcal{H}}$, the identity operator on \mathcal{H} .

A POVM that corresponds to a physical experiment has a simple interpretation. The set Ω is the sample space corresponding to experimental outcomes so that the σ -field \mathcal{F} consists of the set of all events. The POVM \hat{A} and a state ρ on \mathcal{H} induce a measure $\mu_{\rho, \hat{A}}(\cdot) = \text{Tr}\{\rho \hat{A}(\cdot)\}$ on Ω so that $\mu_{\rho, \hat{A}}(E)$ gives the probability of event $E \in \mathcal{F}$. We use the notation $\hat{A} \in E$ to denote the event that the measurement of POVM \hat{A} resulted in a value in $E \in \mathcal{F}$.

Note that here Ω can have any general structure. This allows one to describe measurement apparatuses with outcomes that are physically e.g. multi-dimensional or on a manifold topology like the circle or sphere, which is not possible with standard von Neumann measurements. The latter are indeed equivalent to a special case of POVMs called *Projection Valued Measures (PVMs)*, which require that Ω is a closed subset of the real line and that the range of \hat{A} only consists of commuting projections. The unique correspondence between PVMs (\hat{A}) and self-adjoint operators describing von Neumann measurements (A) is obtained through (1) by setting $P_\omega = \hat{A}(\omega)$ for each $\omega \in \Omega$.

It has been shown that any POVM on \mathcal{H} can be viewed as the restriction to \mathcal{H} of a PVM on a larger Hilbert space.

Theorem 2.1: [6, Th.9.3.2] Let Ω be a compact metrizable space with Borel field \mathcal{F} , and \hat{A} a POVM taking values in $\mathcal{L}_+(\mathcal{H})$. Then there exists a Hilbert space $\mathcal{K} \supset \mathcal{H}$ and a PVM $\hat{A}^{\mathcal{K}} : \mathcal{F} \rightarrow \mathcal{L}_+(\mathcal{K})$ such that if P is the orthogonal projection from \mathcal{K} onto \mathcal{H} then $\hat{A}(E)$ is the restriction of $P \hat{A}^{\mathcal{K}}(E) P$ to \mathcal{H} for all Borel sets E .

In principle, the existing filtering theory for PVMs [11], [12] thus covers the needs of POVM-based filtering, modulo a

²For clarity of explanation, we assume that A has a discrete spectrum. The discussion easily generalizes to the continuous spectrum situation.

proper lift of the Hilbert space. However, the latter is not unique and for reasons explained in Section I, it makes sense to look for a POVM theory that does not build on its reduction from a PVM.

To generalize the second measurement postulate, Davies [6] introduces the notion of an instrument as a complement to POVMs. In the following we denote by $\text{CP}(\mathcal{H})$ the set of all completely positive (CP) maps³ $\mathbf{A} : \rho \mapsto \mathbf{A}(\rho)$ on states $\rho \in \mathcal{S}(\mathcal{H})$.

Definition 2.2: [6, Def 4.1.1] Let Ω be a set, \mathcal{F} a σ -field of subsets of Ω , and \mathcal{H} a Hilbert space. Then a \mathcal{H} -valued instrument on Ω is a map $\hat{\mathbf{A}} : \mathcal{F} \rightarrow \text{CP}(\mathcal{H})$; $E \rightarrow \hat{\mathbf{A}}_E(\cdot)$ such that

- 1) $\hat{\mathbf{A}}_E \geq \hat{\mathbf{A}}_\emptyset = 0$ for all $E \in \mathcal{F}$;
- 2) For any countable, mutually disjoint collection $\{E_n\} \subset \mathcal{F}$ we have

$$\hat{\mathbf{A}}_{(\cup_n E_n)} = \sum_n \hat{\mathbf{A}}_{E_n}$$

where the series convergence on the right is in the weak operator topology;

- 3) $\text{Tr} \{ \hat{\mathbf{A}}_\Omega(\rho) \} = \text{Tr} \{ \rho \}$, for all $\rho \in \mathcal{S}(\mathcal{H})$.

If an experiment is set up so that the outcomes take values in some set Ω , then for a quantum system initially prepared in state $\rho \in \mathcal{S}(\mathcal{H})$ the measurement postulates for an instrument write:

- 1) an outcome in the set $E \subset \Omega$ is obtained with probability $\mathbb{P}(E) = \text{Tr} \{ \hat{\mathbf{A}}_E(\rho) \}$;
- 2) the state of the system conditioned on a measurement outcome in set E is $\hat{\mathbf{A}}_E(\rho) / \mathbb{P}(E)$.

Theorem 2.2: [6, Th.3.1.3, Th.9.2.3] If $\hat{\mathbf{A}}$ is an instrument then for all $E \in \mathcal{F}$, there exists a (non-unique) countable set $\{A_n(E)\}_{n \in \mathbb{N}} \subset \mathcal{L}(\mathcal{H})$ such that

$$\hat{\mathbf{A}}_E(\rho) = \sum_{n \in \mathbb{N}} A_n(E) \rho A_n(E)^* . \quad (2)$$

Moreover, there exists a unique POVM \hat{A} on Ω associated to $\hat{\mathbf{A}}$ such that for all $E \subset \Omega$ and $\rho \in \mathcal{S}(\mathcal{H})$ we have:

$$\text{Tr} \{ \hat{\mathbf{A}}_E(\rho) \} = \text{Tr} \{ \rho \hat{A}_E \} . \quad (3)$$

In particular, this unique POVM is given by

$$\hat{A}(E) = \sum_{n \in \mathbb{N}} A_n(E)^* A_n(E) \text{ for all } E \in \mathcal{F} .$$

Theorem 2.1 implies that it is also always possible to construct an instrument corresponding to a given POVM. However, there is no unique nor canonical way to choose the instrument without further information about the physical system.

³CP maps are the most general possible quantum evolutions, see [6, Sec.9.2] and [17, p.251] for a discussion. The definition given below for instruments is different from that given in [6], wherein the assumption of complete positivity is replaced by positivity; see the discussion in [6, Sec.9.2].

A. System identification of POVMs

Consider an experimental setup corresponding to unknown instrument $\hat{\mathbf{A}}$, associated to POVM \hat{A} . In order to experimentally determine \hat{A} , one can initialize the system being measured in some state $\rho = |\phi\rangle\langle\phi|$ and for any $E \in \mathcal{F}$, the probability of measurement outcome in set E is

$$\mathbb{P}[\hat{A} \in E] = \langle\phi| \hat{A}(E) |\phi\rangle .$$

With sufficiently many experimental outcomes, one can estimate the classical probability distribution $\langle\phi| \hat{A}(E) |\phi\rangle$ over all $E \subset \Omega$; doing this for different $|\phi\rangle$ and using polarization then allows to calculate $\hat{A}(E)$ itself. In order to ascertain the instrument $\hat{\mathbf{A}}$ however, we must have access to the state $\hat{\mathbf{A}}_E(\rho)$ e.g. performing a state tomography experiment. This is often impractical in experimental setups.

In the following sections, we use instruments in theoretical developments and to examine general properties of the measurement settings we consider. The goal is however to show that our final filtering equation works with POVM data only.

B. Notation

In the remainder of the paper we will use $\hat{\mathbf{A}}, \hat{\mathbf{B}}, \dots$ to denote instruments and \hat{A}, \hat{B}, \dots to denote the POVMs corresponding to these instruments. Also, if an instrument $\hat{\mathbf{A}}$ corresponds to a von Neumann measurement, then we denote by A the associated self-adjoint operator. We will use the terms PVM and self-adjoint operator interchangeably.

III. COMMUTING INSTRUMENTS

There is a very general notion of quantum conditional expectation that was first studied in [18] (cf. review in [19]). However, a definition of conditional expectation that is motivated by quantum measurements and is closely related to the classical Kolmogorov definition for random variables is discussed in [11], [12]. The classical definition of conditional expectation builds on joint probabilities, which are not obvious in the quantum context. Therefore the central idea in [11] is that in order to define a conditional expectation of two self-adjoint operators, the two operators must commute with each other. Indeed, in other situations the von Neumann measurement postulates imply that no consistent ‘joint’ measurement can be defined.

We now wish to generalize the [11], [12] definition of conditional expectation for two POVMs and for this we need to understand when it is possible to measure two POVMs simultaneously. In fact this depends on the commutativity of the instruments used to implement the POVMs.

Definition 3.1: Suppose $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ are two measure spaces and \mathcal{H} is some Hilbert space. Then two \mathcal{H} -valued instruments, $\hat{\mathbf{A}}_1 : \mathcal{F}_1 \rightarrow \text{CP}(\mathcal{H})$ and $\hat{\mathbf{A}}_2 : \mathcal{F}_2 \rightarrow \text{CP}(\mathcal{H})$ are said to commute with each other if for all $E_1 \in \mathcal{F}_1$ and $E_2 \in \mathcal{F}_2$, there exist sequences of operators $\{A_n^1(E_1) : n \in N_1\}$ and $\{A_m^2(E_2) : m \in N_2\}$ in $\mathcal{L}(\mathcal{H})$ such

that:

- the instruments write (cf. Theorem 2.2)

$$\hat{\mathbf{A}}_{1,(E_1)}(\rho) = \sum_{n \in N_1} A_n^1(E_1) \rho A_n^1(E_1)^* \text{ for all } \rho \quad (4)$$

$$\hat{\mathbf{A}}_{2,(E_2)}(\rho) = \sum_{n \in N_2} A_n^2(E_2) \rho A_n^2(E_2)^* \text{ for all } \rho; \quad (5)$$

- for all $m \in N_A$ and $n \in N_B$ we have:

$$[A_m^1(E_1), A_n^2(E_2)] = [A_m^1(E_1), A_n^2(E_2)^*] = 0. \quad (6)$$

Remark 3.1: For the special case of PVMs, $N_1 = N_2 = 1$ and the commutativity condition of Definition 3.1 is clearly equivalent to the commutativity of the associated self-adjoint operators A_1 and A_2 . Note that in general, and even when satisfying Definition 3.1, a POVM instrument does not commute with itself (cf. proof of Theorem 4.3.1 in [6]).

Now we consider the composition of instruments — the filtering application will involve one (actual) instrument on the probe and one (hypothetical, expressing our goal-variable) on the target system.

Theorem 3.1: [6, Th.3.4.2] Suppose $\hat{\mathbf{A}}_i$, $i = 1, 2, \dots, n$ are instruments on some compact metrizable Ω_i with Borel field \mathcal{F}_i . Then there exists a unique “joint” instrument $\hat{\mathbf{A}}$ on $\Omega_1 \times \Omega_2 \times \dots \times \Omega_n$ such that for all $E_i \in \mathcal{F}_i$ and ρ we have:

$$\hat{\mathbf{A}}_{E_1 \times E_2 \times \dots \times E_n}(\rho) = \hat{\mathbf{A}}_{n,(E_n)} \circ \dots \circ \hat{\mathbf{A}}_{2,(E_2)} \circ \hat{\mathbf{A}}_{1,(E_1)}(\rho).$$

We now prove the first result of this paper.

Theorem 3.2: Suppose $\hat{\mathbf{A}}_i$, $i = 1, 2, \dots, n$ are instruments on some compact metrizable Ω_i with Borel field \mathcal{F}_i and \hat{A}_i are the corresponding POVMs. If the \hat{A}_i are mutually commutative, then the POVM \hat{A} corresponding to the joint instrument $\hat{\mathbf{A}}$ is uniquely determined by the POVMs \hat{A}_i , according to

$$\hat{A}(E_1 \times E_2 \times \dots \times E_p) = \hat{A}_1(E_1) \hat{A}_2(E_2) \dots \hat{A}_p(E_p).$$

Moreover, the POVMs are mutually commutative, that is $[\hat{A}_i(E_i), \hat{A}_j(E_j)] = 0$ for all E_i, E_j , $i \neq j$.

Proof: We first prove the result for $p = 2$. From Definition 3.1 construct sequences of operators $\{A_n^1 : n \in N_A\}$ and $\{A_n^2 : n \in N_B\}$ satisfying (4),(5),(6). Select some events $E_1 \in \mathcal{F}_1$, $E_2 \in \mathcal{F}_2$ and let $E = E_1 \times E_2$. Then for all $\rho \in \mathcal{S}(\mathcal{H})$, we have

$$\begin{aligned} \text{Tr} \left\{ \rho \hat{A}_E \right\} &= \text{Tr} \left\{ \hat{\mathbf{A}}_E(\rho) \right\} \\ (\text{Th.3.1}) &= \text{Tr} \left\{ \hat{\mathbf{A}}_{2,(E_2)} \circ \hat{\mathbf{A}}_{1,(E_1)}(\rho) \right\} \\ (\text{Th.2.2}) &= \text{Tr} \left\{ \sum_{m \in N_1, n \in N_2} A_n^2 A_m^1 \rho A_m^1{}^* A_n^2{}^* \right\} \\ (\text{trace property}) &= \text{Tr} \left\{ \rho \sum_{m \in N_1, n \in N_2} A_m^1{}^* A_n^2{}^* A_n^2 A_m^1 \right\} \\ (\text{Def.3.1}) &= \text{Tr} \left\{ \rho \sum_{m \in N_1} A_m^1{}^* A_m^1 \sum_{n \in N_2} A_n^2{}^* A_n^2 \right\} \\ &= \text{Tr} \left\{ \rho \hat{A}_1 \hat{A}_2 \right\}, \end{aligned} \quad (7)$$

where for notational convenience, we have written A_n^i for $A_n^i(E_i)$, with $i = 1, 2$. Taking the adjoint of the trace

argument in the last expression and cycling ρ to the left finally yields $\text{Tr} \left\{ \rho \hat{A}_1 \hat{A}_2 \right\} = \text{Tr} \left\{ \rho \hat{A}_2 \hat{A}_1 \right\}$.

When adding a third instrument $\hat{\mathbf{A}}_3$ to this setting, we assume that $\hat{\mathbf{A}}_3$ commutes with $\hat{\mathbf{A}}_1$ and $\hat{\mathbf{A}}_2$, but a priori the representations (2) used for these various commutations might be incompatible. We now show that the developments above can nevertheless be repeated. From Theorem 3.1 we know that $\hat{\mathbf{A}}$ is uniquely determined by $\hat{\mathbf{A}}_1$ and $\hat{\mathbf{A}}_2$, and by Theorem 2.2 the POVM, \hat{A} is uniquely determined by $\hat{\mathbf{A}}$. Therefore, given any two sequences of operators $\{B_n^1 : n \in M_1\}$ and $\{B_n^2 : n \in M_2\}$ such that

$$\hat{\mathbf{A}}_{i,(E_i)}(\rho) = \sum_{n \in M_i} B_n^i \rho B_n^{i*} \quad (8)$$

for $i = 1, 2$ — here, we might choose $\{B_n^1\}$ that do not commute with the $\{B_n^2\}$ — we have

$$\sum_{n \in N_i} A_n^i{}^* A_n^i = \sum_{n \in M_i} B_n^{i*} B_n^i$$

and also

$$\begin{aligned} \text{Tr} \left\{ \hat{\mathbf{A}}_E(\rho) \right\} &= \text{Tr} \left\{ \hat{\mathbf{A}}_{1,(E_1)} \circ \hat{\mathbf{A}}_{2,(E_2)}(\rho) \right\} \\ (\text{definition}) &= \text{Tr} \left\{ \sum_{m \in M_1, n \in M_2} B_m^1 B_n^2 \rho B_n^{2*} B_m^{1*} \right\} \\ (\text{put (8) in (7)}) &= \text{Tr} \left\{ \rho \sum_{n \in M_2} B_n^{2*} B_n^2 \sum_{m \in M_1} B_m^1{}^* B_m^1 \right\}. \end{aligned}$$

This allows to prove the theorem for $p > 2$ with a recursive argument. \blacksquare

IV. CONDITIONAL EXPECTATION

We are now in a position to define the conditional expectation of POVMs associated with two commuting instruments. Recall that we wish to find an expression for the conditional expectation that is expressed only in terms of the POVMs and not in terms of the instruments themselves.

A. Basic definitions

In this Section, we assume that \mathcal{H} is a Hilbert space, Ω_A and Ω_B are two compact metrizable sets with Borel algebras \mathcal{F}_A and \mathcal{F}_B , respectively, and \hat{A} and \hat{B} are two \mathcal{H} -valued POVMs on Ω_A and Ω_B corresponding to the instruments $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$.

Suppose ρ is a state on \mathcal{H} and f is any measurable function on Ω_A . Then the integral of f with respect to \hat{A} over a set $E \in \mathcal{F}_A$ is defined by:⁴

$$\int_{\omega \in E} f(\omega) \text{Tr} \left\{ \rho d\hat{A}(\omega) \right\} \triangleq \int_{\omega \in E} f(\omega) d\mu_{\rho, \hat{A}}(\omega)$$

where the measure $\mu_{\rho, \hat{A}}(\cdot) \triangleq \text{Tr} \left\{ \rho \hat{A}(\cdot) \right\}$ is a probability measure on $(\Omega_A, \mathcal{F}_A)$. When a vector space is associated

⁴It should also be possible to define this integral as a Stieltjes integral over the measurable space $\mathcal{L}_+(\mathcal{H})$ with the measure induced by the POVM.

to the POVM measurement results Ω_A , we can compute the expectation value of a POVM \hat{A} in state ρ by:

$$\mathbb{E}_\rho[\hat{A}] = \int_{\omega \in \Omega} \omega \text{Tr} \left\{ \rho d\hat{A}(\omega) \right\}.$$

Defining the inverse map $\hat{A}^{-1} : \mathcal{L}_+(\mathcal{H}) \rightarrow \Omega_A$ of \hat{A} in the standard way, we can rewrite this as

$$\mathbb{E}_\rho[\hat{A}] = \int_{X \in \hat{A}(\Omega)} \hat{A}^{-1}(X) \text{Tr} \left\{ \rho dX \right\}. \quad (9)$$

Now suppose \hat{A} and \hat{B} are such that $\hat{A}(\mathcal{F}_A) \subset \hat{B}(\mathcal{F}_B)$. Then for any $E \in \mathcal{F}_A$ we define

$$\begin{aligned} & \int_{\omega \in E} \left(\hat{B}^{-1} \hat{A}(\omega) \right) \text{Tr} \left\{ \rho d\hat{A}(\omega) \right\} \\ & \triangleq \int_{\omega \in E} \left(\hat{B}^{-1} \hat{A}(\omega) \right) d\mu_{\rho, \hat{A}}(\omega) \\ & \triangleq \lim_{P_n \rightarrow 0} \sum_n \bar{\omega}_n \mu_{\rho, \hat{A}}(P_n) \end{aligned}$$

if the limit exists. Here $\{P_n\}$ is some partition of Ω_A and $\bar{\omega}_n$ is in the smallest set $\bar{P}_n \in \mathcal{F}_B$ such that $\hat{B}(\bar{P}_n) \geq \hat{A}(P_n)$.

B. Main result

For a given state ρ and two commuting instruments $\hat{\mathbf{A}}, \hat{\mathbf{B}}$ with Ω_A a convex subset of a vector space, in analogy with the classical definition we would like to define the conditional expectation $\mathbb{E}[\hat{A}|\hat{B}] : \Omega_A \times \Omega_B \rightarrow \Omega_A$ of \hat{A} with respect to \hat{B} such that $\mathbb{E}[\hat{A}|\hat{B}]$ is \mathcal{F}_B -measurable and:

$$\begin{aligned} & \int_{\omega=(\omega_A, \omega_B) \in \Omega_A \times E} \omega_A \text{Tr} \left\{ \rho d(\hat{A}\hat{B})(\omega) \right\} \\ & = \int_{\omega=(\omega_A, \omega_B) \in \Omega_A \times E} \mathbb{E}[\hat{A}|\hat{B}](\omega) \text{Tr} \left\{ \rho d(\hat{A}\hat{B})(\omega) \right\} \end{aligned} \quad (10)$$

for all $E \in \mathcal{F}_B$. Moreover, referring to (9), we want to define a POVM $\hat{\mathbb{E}}_{\hat{A}|\hat{B}} : \Omega_A \rightarrow \mathcal{L}_+(\mathcal{H})$ associated to this conditional expectation, such that

$$\begin{aligned} & \int_{\omega=(\omega_A, \omega_B) \in \Omega_A \times E} \omega_A \text{Tr} \left\{ \rho d(\hat{A}\hat{B})(\omega) \right\} \\ & = \int_{\omega_B \in E} \left(\hat{\mathbb{E}}_{\hat{A}|\hat{B}}^{-1} \hat{B}(\omega_B) \right) \text{Tr} \left\{ \rho d\hat{B}(\omega_B) \right\} \end{aligned} \quad (11)$$

for all $E \in \mathcal{F}_B$. The following proves that this is possible.

Theorem 4.1: Suppose ρ is a state on \mathcal{H} and $\hat{\mathbf{A}}, \hat{\mathbf{B}}$ are commuting instruments with Ω_A a convex subset of a vector space. Then there exists a σ -field \mathcal{G}_A on Ω_A and a (ρ -unique) POVM $\hat{\mathbb{E}}_{\hat{A}|\hat{B}} : \mathcal{G}_A \rightarrow \mathcal{L}_+(\mathcal{H})$ satisfying (11). We call $\hat{\mathbb{E}}_{\hat{A}|\hat{B}}$ the *conditional POVM for \hat{A} with respect to \hat{B}* .

Proof: From Theorems 3.1 and 3.2 we know that the POVM $\hat{A}\hat{B}$ on $\Omega \triangleq \Omega_A \times \Omega_B$ is given by $\hat{A}\hat{B}(E_A \times E_B) = \hat{A}(E_A)\hat{B}(E_B)$. Denote by $\mu = \mu_{\rho, \hat{A}\hat{B}}(\cdot)$ the measure induced by $\hat{A}\hat{B}$ on Ω .

We now define two classical random variables

$$\alpha : (\omega_A, \omega_B) \mapsto \omega_A \text{ and } \beta : (\omega_A, \omega_B) \mapsto \omega_B$$

on Ω . Then we know from the classical Kolmogorov theory that the conditional expectation

$$\mathbb{E}[\alpha|\beta] : \Omega \rightarrow \Omega_A$$

exists, is a (μ -a.s.) unique random variable that is measurable with respect to the σ -algebra \mathcal{F}_β generated by β and satisfies for all $E \in \mathcal{F}_\beta$,

$$\int_{\omega \in E} \alpha(\omega) d\mu(\omega) = \int_{\omega \in E} \mathbb{E}[\alpha|\beta] d\mu(\omega). \quad (12)$$

Set \mathcal{G}_A to be the σ -algebra generated by $\mathbb{E}[\alpha|\beta]$ in Ω_A . Clearly, $\mathcal{F}_\beta = \mathcal{F}_B$ and for all $E \in \mathcal{G}_A$ the \mathcal{F}_β -measurability of $\mathbb{E}[\alpha|\beta]$ implies that $\mathbb{E}[\alpha|\beta]^{-1}(E) \in \mathcal{F}_B$. Therefore, we can define $\hat{\mathbb{E}}_{\hat{A}|\hat{B}}$ as follows: for any $E \in \mathcal{G}_A$ set

$$\hat{\mathbb{E}}_{\hat{A}|\hat{B}}(E) \triangleq \hat{B} \left(\mathbb{E}[\alpha|\beta]^{-1}(E) \right). \quad (13)$$

Now (11) is simply a restatement of (12). Also, the ρ -uniqueness of $\hat{\mathbb{E}}_{\hat{A}|\hat{B}}$ follows from that of the classical conditional expectation $\mathbb{E}[\alpha|\beta]$. ■

C. Example

We now study a simple example to understand the ideas developed above. We note that it is possible to solve this example problem quite easily using some simple algebraic methods. However, the main utility of our novel formulation of conditional expectation becomes apparent in the case where the sample spaces of the POVMs are continuous valued as in the case of phase measurements [20].

In this example we assume that \mathcal{H}_s is a finite dimensional Hilbert space \mathbb{C}^d . We further assume that the system interacts with a single probe with state space \mathcal{H}_p and the system and probe are initially in the separable state $\rho = \rho_s \otimes \rho_p$ in $\mathcal{H}_{\text{tot}} = \mathcal{H}_s \otimes \mathcal{H}_p$, where the initial state of the probe is a pure state $\rho_p = |\psi_p\rangle\langle\psi_p|$. The system and probe undergo a unitary evolution U , after which a \mathcal{H}_p -valued POVM \hat{B} on Ω_B , corresponding to an instrument $\hat{\mathbf{B}}$, is measured on the probe. We assume that \hat{B} is discrete, that is there exists a discrete set $\{\omega_1, \omega_2, \dots\} \subset \Omega_B$ such that $\hat{B}(\{\omega_1, \omega_2, \dots\}) = \mathbb{I}_{\mathcal{H}_p}$. Let \hat{A} be a discrete \mathcal{H}_s -valued PVM on $\Omega_A = \mathbb{R}$, associated to the self-adjoint operator A on \mathcal{H}_s . Note that we have to associate Ω_A to a vector space, but we make no such assumption on Ω_B i.e. for the measurement carried out on the probe. This will allow us later to consider probe measurements with results on manifolds, like the circle for phase measurements.

For any \mathcal{H}_{tot} valued POVM \hat{W} we denote by \hat{W}^U the POVM $U^* \hat{W} U$ on the same measure space as \hat{W} obtained from the unitary evolution U on \hat{W} (Heisenberg picture). We can extend any POVM defined on \mathcal{H}_s or \mathcal{H}_p to the tensor product $\mathcal{H}_{\text{tot}} = \mathcal{H}_s \otimes \mathcal{H}_p$; making a slight abuse of notation, we use the same symbol for the POVM on the tensor product space. Clearly, any reasonable instruments associated to the \mathcal{H}_{tot} -valued POVMs \hat{A} and \hat{B} should commute, because those measurements are physically implemented on different tensor factors of \mathcal{H}_{tot} . Therefore, the instruments corresponding to the POVMs \hat{A}^U and \hat{B}^U also commute. Then the conditional POVM $\hat{\mathbb{E}}_{\hat{A}^U|\hat{B}^U}$ is well defined and the map

$$\Phi : A \mapsto \int_{\omega_B \in E} \left([\hat{\mathbb{E}}_{\hat{A}^U|\hat{B}^U}]^{-1} \hat{B}^U(\omega_B) \right) \text{Tr} \left\{ \rho d\hat{B}^U(\omega_B) \right\}$$

defines a real valued linear functional on the self-adjoint operators on \mathcal{H}_s that gives the expectation value of observable

A given measurement outcome $\hat{B} \in E$. Therefore, there exists a density operator $\hat{\rho}_s$ on \mathcal{H}_s , which we identify with the post-measurement state of the system, such that

$$\text{Tr}\{A\hat{\rho}_s\} = \Phi(A).$$

Using Equation (13) with the notation of the classical random variables α and β from the proof of Theorem 4.1, we have $\Phi(A) = \int_{\omega \in E} \mathbb{E}[\alpha|\beta] d\mu(\omega)$, and for an elementary event $E = \{\omega_r\} \in \Omega_B$ some algebra gives us

$$\text{Tr}\{A\hat{\rho}_s\} = \frac{\text{Tr}\left\{[\rho_s \otimes \rho_p] \hat{B}^U(\omega_r) \left(\sum_{\omega_a \in \Omega_A} \omega_a \hat{A}^U(\omega_a)\right)\right\}}{\text{Tr}\left\{[\rho_s \otimes \rho_p] \hat{B}^U(\omega_r)\right\}}.$$

Using the definition of $(\cdot)^U$, the cyclic property of the trace, and the fact that $\text{Tr}\{\rho A\} = \text{Tr}\left\{\sum_{\omega_a \in \Omega_A} \omega_a \hat{A}(\omega_a)\right\}$ for a self-adjoint operator A , where Ω_A is the spectrum of A , we get

$$\text{Tr}\{A\hat{\rho}_s\} = \frac{\text{Tr}\left\{U[\rho_s \otimes \rho_p]U^* \hat{B}(\omega_r) A\right\}}{\text{Tr}\left\{U[\rho_s \otimes \rho_p]U^* \hat{B}(\omega_r)\right\}}.$$

Because this is true for all A we have

$$\hat{\rho}_s = \frac{\text{Tr}_p\left\{U[\rho_s \otimes \rho_p]U^* \hat{B}(\omega_r)\right\}}{\text{Tr}\left\{U[\rho_s \otimes \rho_p]U^* \hat{B}(\omega_r)\right\}}. \quad (14)$$

V. APPLICATION TO QUANTUM FILTERING

A. Measurement model

Our measurement model is motivated from the discrete-time model used in [12] for filtering using PVMs. We consider a system with Hilbert space \mathcal{H}_s and a probe consisting of a sequence of subsystems $n = 1, 2, \dots$ each with Hilbert space $\mathcal{H}_n = \mathcal{H}$. So the probe is described on a Hilbert space $\mathcal{H}_p = \otimes_{n=1}^{\infty} \mathcal{H}_n$, the combined state space of the probe and system is written $\mathcal{H}_{tot} = \mathcal{H}_s \otimes \mathcal{H}_p$. Suppose Ω_n is a compact metrizable space⁵ for $n = 1, 2, \dots$, let \mathcal{F}_n a σ -field on Ω_n and $\hat{B}_n : \mathcal{F}_n \rightarrow \text{CP}(\mathcal{H}_n)$ a \mathcal{H}_n valued instrument with corresponding POVM \hat{B}_n . Also, let $(\Omega_A, \mathcal{F}_A)$ be some measure space and $\hat{A} : \Omega_A \rightarrow \mathcal{H}_s$ any system instrument with corresponding POVM \hat{A} . We set $\mathcal{H}_n = \mathcal{H}_s$ for $n = 0$ and recursively define $\mathcal{H}_n = \mathcal{H}_{n-1} \otimes \mathcal{H}_n$ for $n \geq 1$; similarly define $\Omega_n = \Omega_A \times \Omega_1 \times \dots \times \Omega_n$. Also set $\mathcal{H}_{(n)} = \otimes_{i=n+1}^{\infty} \mathcal{H}_i$ so that $\mathcal{H}_{tot} = \mathcal{H}_n \otimes \mathcal{H}_{(n)}$.

Consider the initial state $\rho_0^{tot} = |\psi_0^{tot}\rangle \langle \psi_0^{tot}|$ on \mathcal{H}_{tot} , with

$$|\psi_0^{tot}\rangle = |\psi_0^s\rangle \otimes_{n=1}^{\infty} |\psi_n^p\rangle.$$

Thus $|\psi^s\rangle$ is the initial system state and each $|\psi_n^p\rangle$ is the initial state of a probe subsystem on \mathcal{H}_n .

We suppose that between time steps n and $n+1$, the system interacts with the probe according to a unitary evolution operator U_n on $\mathcal{H} \otimes \mathcal{H}_n$, i.e. it interacts only with subsystem n of the probe. After this unitary evolution, the POVM \hat{B}_n is measured to be some $\omega_n \in \Omega_n$. We wish to find a recursive equation for ρ_n conditioned on the measurement outcome $\hat{B}_n = \omega_n$.

⁵If Ω_n is not compact then we can simply consider the 1-point compactification of Ω_n [6, p.12].

B. Filtering equation

With a slight abuse of notation, let $U_n = \prod_{i=1}^n U_i$, a unitary on \mathcal{H}_n . As in Section IV-C, instruments \hat{A} and \hat{B}_i commute for all i because they are defined on different Hilbert spaces in the tensor product space \mathcal{H}_{tot} , and clearly the same applies to instruments \hat{B}_i, \hat{B}_j for all $i \neq j$. Then the same commutations apply to the evolved instruments

$$\hat{A}(n) \triangleq U_n \hat{A} U_n^* \quad \text{and} \quad \hat{B}_i(n) \triangleq U_n \hat{B}_i U_n^*.$$

Therefore, the conditional expectation $\mathbb{E}[\hat{A}(n)|\hat{B}_1(n)\hat{B}_2(n)\dots\hat{B}_n(n)]$ is a well defined function of the measurement outcomes $\omega_1, \dots, \omega_n$, and also the associated POVM

$$\hat{\mathbb{E}}_{\hat{A}(n)|\hat{B}_1(n)\hat{B}_2(n)\dots\hat{B}_n(n)}$$

is well-defined. The main advantage of the commutativity condition is seen at this point: through the conditional expectation POVM, the conditional expectation of a system POVM \hat{A} is an explicit function of the measurement outcomes if the instruments $\hat{A}, \hat{B}_1, \dots, \hat{B}_n$ are mutually commutative.

Now we specialize to the case where the system instrument \hat{A} corresponds to a self-adjoint operator $A : \mathcal{H}_s \rightarrow \mathcal{H}_s$, i.e. the instrument \hat{A} corresponds to a PVM. The linearity of the conditional expectation based on $\hat{\mathbb{E}}_{\hat{A}(n)|\hat{B}_1(n)\hat{B}_2(n)\dots\hat{B}_n(n)}$ implies that

$$\Psi : A \mapsto \int_{\omega \in \Omega_A} \omega \text{Tr} \left\{ \rho_0^{tot} d\hat{\mathbb{E}}_{\hat{A}(n)|\hat{B}_1(n)\hat{B}_2(n)\dots\hat{B}_n(n)} \right\}$$

is a real valued linear map on the set of self-adjoint operators on \mathcal{H}_s . Therefore there exists a unique density operator ρ_n^s on \mathcal{H}_s that satisfies

$$\Psi(A) = \text{Tr}\{\rho_n A\}$$

for all self-adjoint $A : \mathcal{H} \rightarrow \mathcal{H}$. Therefore, ρ_n is the estimated state of the system at time step n . Explicit expressions like (14) could be developed along the same lines as in Section IV-C.

VI. CONCLUSION AND FUTURE WORK

In this paper we show how quantum filtering can be performed in the POVM setting. We therefore formalize a notion of commuting instruments for two measurements, which gives a sufficient condition to define joint measurement statistics in terms of the associated POVMs only, without explicitly depending on the instruments. We then introduce the notion of conditional-expectation-POVM for two measurements with commuting instruments. On that basis, for a system-probe model, we establish the filtering equation for the system state conditioned on probe measurements with a general instrument, and we demonstrate that it is only a function of the probe POVMs. In a sequel to this paper [20], we will apply the developed ideas to derive a filtering equation for discrete-time POVM phase measurements, which is a quintessential example of a POVM-type measurement where the associated instrument is not known. On a more general mode, we also intend to explore *necessary* conditions for the commutation of two instruments.

REFERENCES

- [1] A. Bensoussan, *Stochastic Control of Partially Observable Systems*. Cambridge University Press, 1992.
- [2] V. P. Belavkin, "Quantum filtering of markov signals with white quantum noise," *Radiotekhnika i Elektronika*, vol. 25, pp. 1445–1453, 1980.
- [3] —, "Theory of the control of observable quantum-systems," *Automation and Remote Control*, vol. 44, p. 178, 1983.
- [4] —, *Nondemolition stochastic calculus in fock space and nonlinear filtering and control in quantum systems*, ser. Stochastic methods in mathematics and physics. XXIV Karpacz winter school: World Scientific, Singapore, 1988, pp. 310–324.
- [5] —, "Quantum stochastic calculus and quantum nonlinear filtering," *Journal of Multivariate Analysis*, vol. 42, no. 2, pp. 171 – 201, 1992. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0047259X9290042E>
- [6] E. B. Davies, *Quantum theory of open systems*. IMA, 1976.
- [7] V. B. Braginsky and F. Y. Khalili, *Quantum Measurement*. Cambridge University Press, 1992.
- [8] S. Haroche and J. Raimond, *Exploring The Quantum: Atoms, Cavities, And Photons (oxford Graduate Texts)*. Oxford University Press, 2006.
- [9] C. W. Gardiner and P. Zoller, *Quantum noise: a handbook of Markovian and non-Markovian quantum stochastic methods with applications to quantum optics*. Springer, 2004, vol. 56.
- [10] H. M. Wiseman and G. J. Milburn, *Quantum measurement and control*. Cambridge University Press, 2010.
- [11] L. Bouten, R. Van Handel, and M. James, "An introduction to quantum filtering," *SIAM Journal on Control and Optimization*, vol. 46, no. 6, pp. 2199–2241, 2007. [Online]. Available: <http://epubs.siam.org/doi/abs/10.1137/060651239>
- [12] L. Bouten, R. Van Handel, and M. R. James, "A discrete invitation to quantum filtering and feedback control," *SIAM review*, vol. 51, no. 2, pp. 239–316, 2009.
- [13] J. H. Shapiro and S. R. Shepard, "Quantum phase measurement: A system-theory perspective," *Phys. Rev. A*, vol. 43, pp. 3795–3818, Apr 1991. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevA.43.3795>
- [14] M. Nielsen and I. Chuang, *Quantum computation and quantum information*. Cambridge university press, 2010.
- [15] D. Berry, "Adaptive Phase Measurements," Ph.D. dissertation, University of Queensland, Feb. 2002.
- [16] R. Lynch, "The quantum phase problem: a critical review," *Physics Reports*, vol. 256, no. 6, pp. 367 – 436, 1995. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/037015739400095K>
- [17] K. Parthasarathy, *An introduction to quantum stochastic calculus*. Springer, 1992.
- [18] H. Umegaki, "Conditional expectation in an operator algebra, i," *Tohoku Mathematical Journal*, vol. 6, no. 2-3, pp. 177–181, 1954.
- [19] B. Kümmerer, "Quantum markov processes," *Coherent evolution in noisy environments*, pp. 139–198, 2002.
- [20] R. Somaraju, A. Sarlette, and H. Thienpont, "Discrete time filtering using POVM measurements: Phase measurements," 2013, in preparation.